

FIG. 2

URL TABLE 120

DOCUMENT ID	URL	ACCUMULATE FLAG	COMMUNITY FLAG	SIGNIFICANCE LEVEL
doc1	http://www.abc.com/	1	1
doc2	http://www.klm.com/	1	0
....

FIG. 3

REFERENCE TABLE 121

REFERRING DOCUMENT ID	REFERENCED DOCUMENT ID ₁ (INSIDE COMMUNITY)	REFERENCED DOCUMENT ID ₂ (OUTSIDE COMMUNITY)
doc1	doc2, doc3, doc4	doc5, doc6, doc7
doc2	doc4, doc8	doc9, doc10
....

FIG. 4

REFERENCING EXPRESSION TABLE 122

EXPRESSION ID (w)	EXPRESSION CHARACTER STRING	FREQUENCY (OF (W))	NECESSITY/ NON-NECESSITY FLAG
r w 1	ABC COMPANY	2 4	1
r w 2	RETURN TO TOP	2 8 3 0 5	0
.....

FIG. 5

COLLECTED DOCUMENT FILE 123

EXPRESSION ID(w) DOCUMENT ID(d)	rw1	rw2	...	rwj	...	rw m
doc1	19	0	
doc2	0	2
...						
doci				TF(doc i, rw j)		
...						
doon

FIG. 6

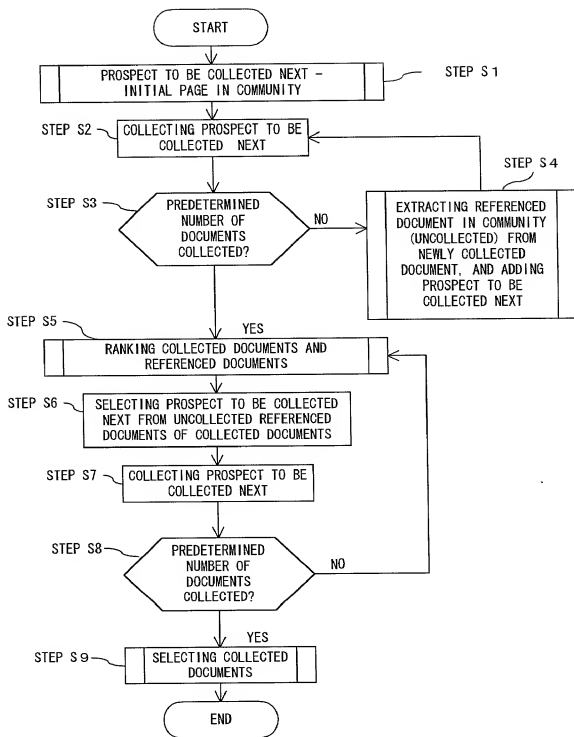


FIG. 7

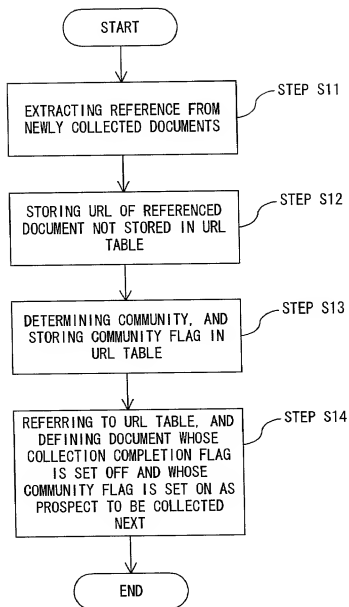


FIG. 8

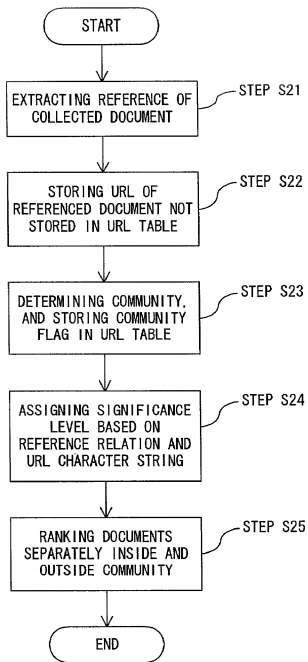


FIG. 9

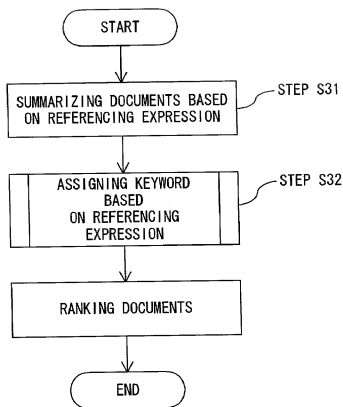


FIG. 10

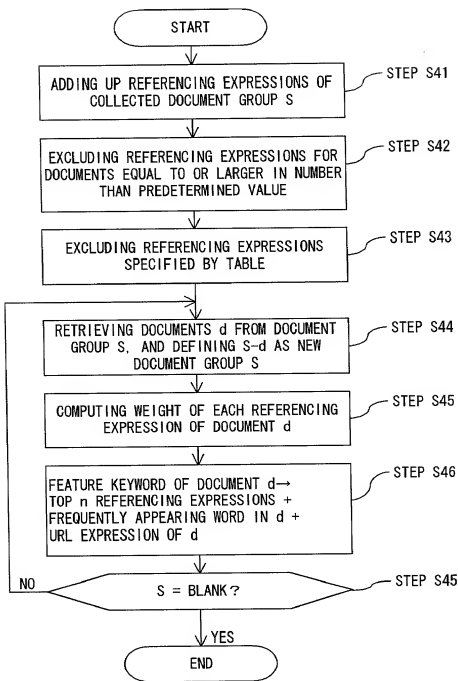


FIG. 11

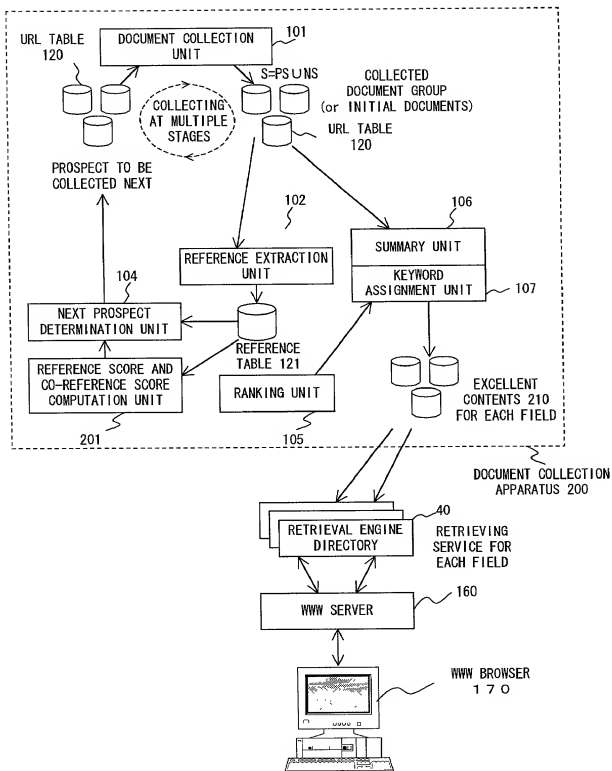


FIG. 13

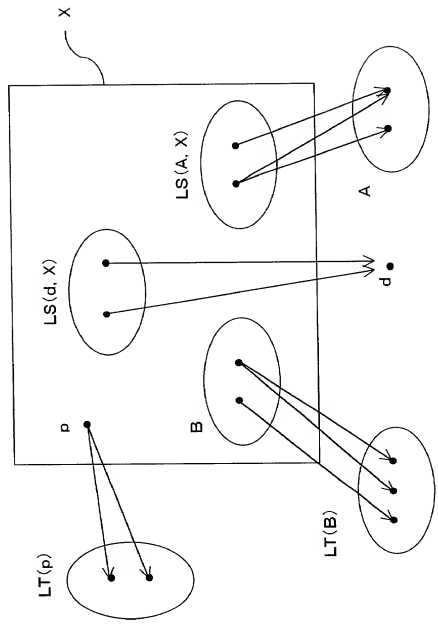


FIG. 14

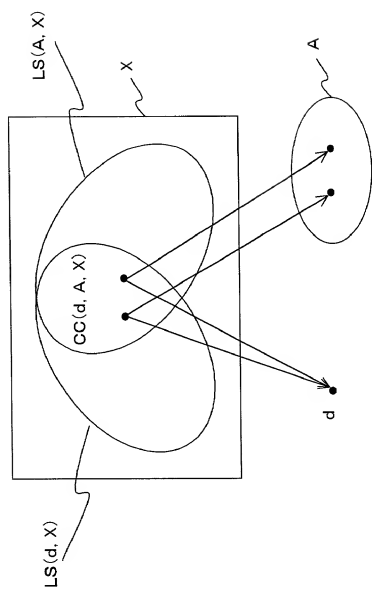


FIG. 15

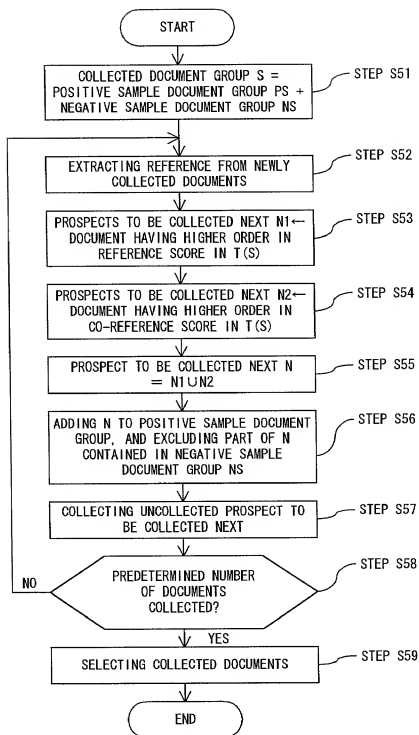


FIG. 16

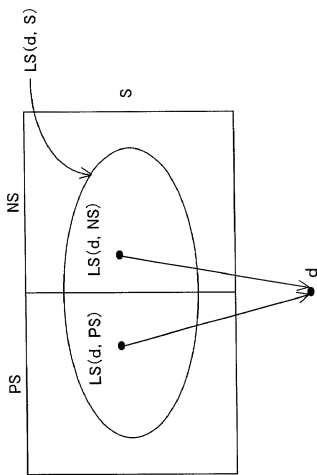


FIG. 17

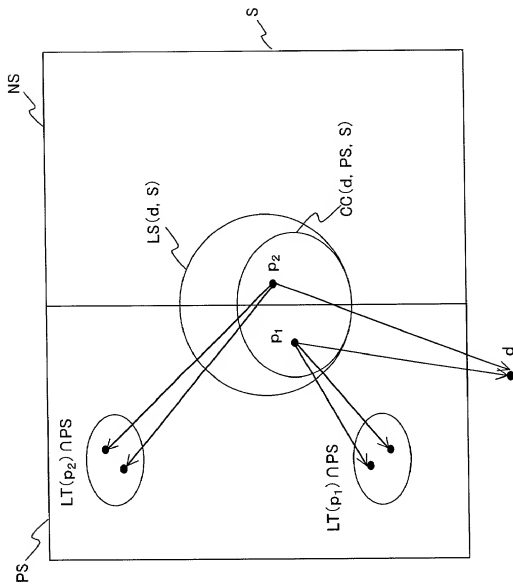


FIG. 18

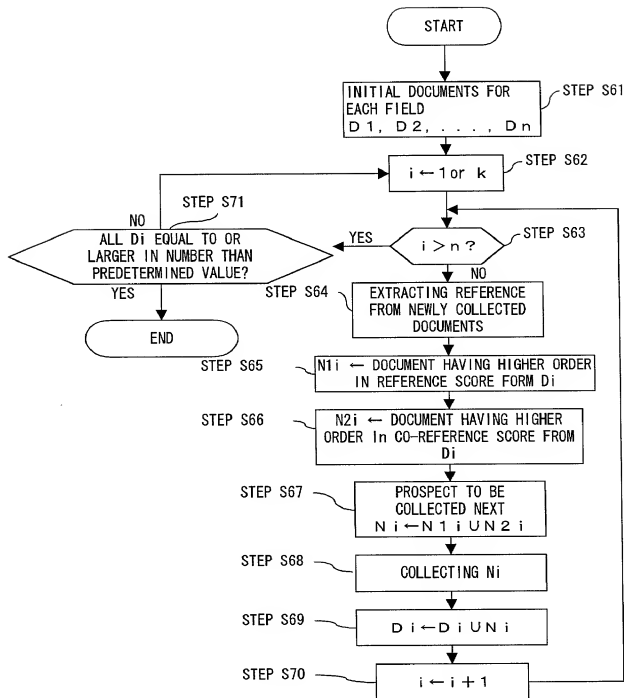


FIG. 19

precision/recall (PRECISION FACTOR /
RECALL FACTOR)

LINUX (CV3. 150000 / 6.73 million)

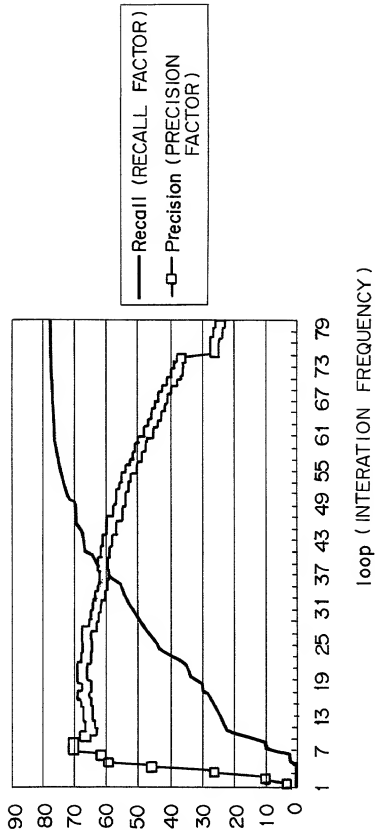


FIG. 20

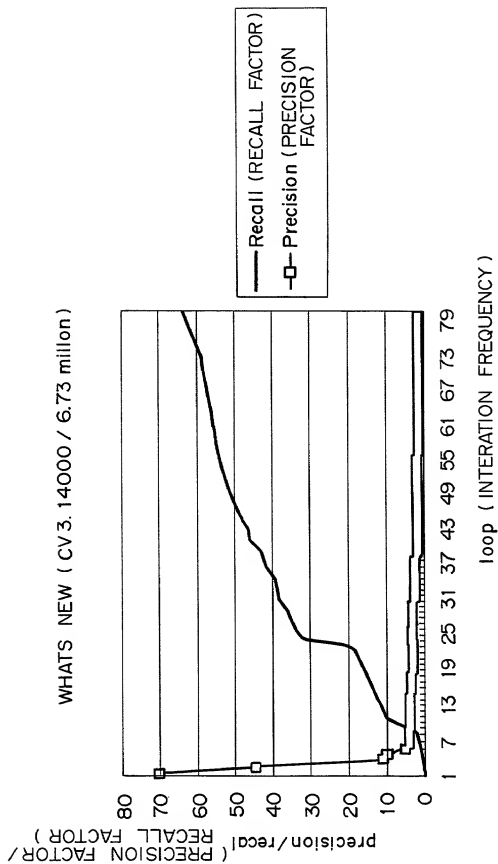
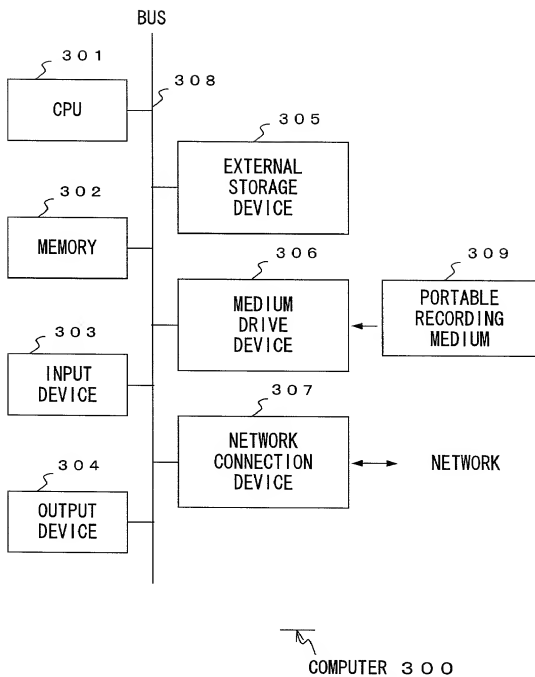


FIG. 21



F I G. 2 2

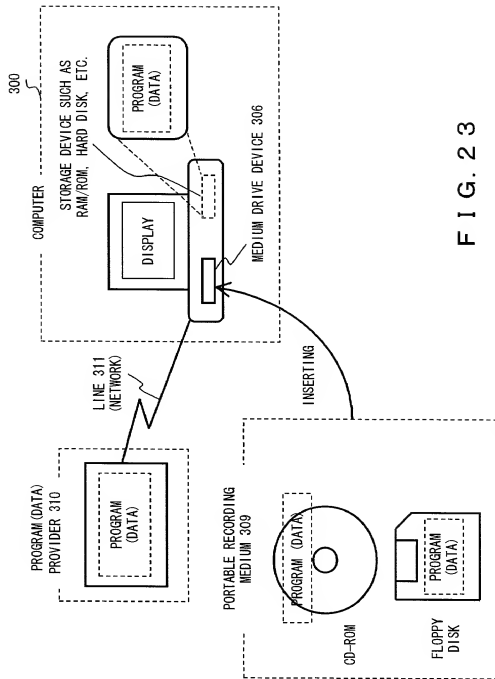


FIG. 23